

Optimal design for EEG/MEG source analysis

H.M. Huizenga, D.J. Heslenfeld, and P.C.M. Molenaar

Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

1 Introduction

The precision of source estimates is dependent on sensor positioning. Optimal positions can be found by minimizing the generalized variance of source parameters as a function of sensor positions. In this paper we describe this so-called optimal design technique [1]. In addition, we report a simulation study on its validity and merits, and we report an empirical illustration.

2 Methods

2.1 Source analysis

The regression model is $\mathbf{y} = \mathbf{f}(\mathbf{X}, \boldsymbol{\theta}) + \boldsymbol{\epsilon}$. The vector $\mathbf{y}$ contains potentials measured on N sensors. $\mathbf{f}(\mathbf{X}, \boldsymbol{\theta})$ is the vector with modeled potentials. $\mathbf{X}$ is the matrix with sensor coordinates, $\boldsymbol{\theta}$ the vector with P source parameters of D sources, and $\boldsymbol{\epsilon}$ the noise vector.

Let $\hat{\mathbf{G}}$ be the N by P matrix with the first order partial derivatives of the N modeled potentials to the P parameters. The covariance matrix of the parameter estimates is then [2]:

$$\hat{\mathbf{C}} = \hat{\sigma}^2 [\hat{\mathbf{G}}' \hat{\mathbf{G}}]^{-1}$$

where $\hat{\sigma}^2$ is an estimate of the noise variance.

A test on the source parameters can be derived in the following way. Let the vector $\mathbf{r}(\hat{\boldsymbol{\theta}})$ contain Q (non)linear hypotheses on $\hat{\boldsymbol{\theta}}$ and let the Q by P matrix $\mathbf{R}(\hat{\boldsymbol{\theta}})$ contain first order partial derivatives of the Q hypotheses to the P parameter estimates. Then

$$[\mathbf{r}(\hat{\boldsymbol{\theta}}) - \mathbf{q}]' [\mathbf{R}(\hat{\boldsymbol{\theta}}) \hat{\mathbf{C}} \mathbf{R}(\hat{\boldsymbol{\theta}})']^{-1} [\mathbf{r}(\hat{\boldsymbol{\theta}}) - \mathbf{q}] \sim Q F_{Q, N-P}^{\alpha}$$

the test statistic for $\mathbf{r}(\hat{\boldsymbol{\theta}}) - \mathbf{q} = \mathbf{0}$ is [2]:

where F denotes the $1-\alpha$ quantile of the F distribution with Q and $N-P$ degrees of freedom.

2.2 Increasing accuracy

The covariance matrix of the parameter estimates depends on $\mathbf{G}$, and thus on the sensor positions $\mathbf{X}$,

and the source parameters $\hat{\boldsymbol{\theta}}$. Our general objective is to determine the position of sensors required for a covariance matrix with the desired properties. Since the covariance matrix depends on $\hat{\boldsymbol{\theta}}$, the method should be based on an explicit source hypothesis. This hypothesis can be derived from anatomical knowledge, from animal studies, or from PET or fMRI studies. It can also be derived from previously reported EEG/MEG studies, or from a pilot study, as will be illustrated later.

The sensor configuration generating an optimal covariance matrix is determined by minimizing the generalized variance of the parameters, as indexed by the determinant $|\mathbf{C}|$ [1]. Equivalently, $|\mathbf{G}'\mathbf{G}|$ can be maximized since σ^2 is independent of sensors under the assumption of homoscedastic and uncorrelated noise. The actual optimization of this so-called D-optimality criterion will be treated later. After the optimal sensor configuration is calculated, it should be determined whether this sensor configuration yields the required accuracy. This can be assessed by computing confidence regions given the source hypothesis and given an initial guess of the noise variance. Therefore these regions will be called a-priori confidence regions. Alternatively, if an optimal design is computed for several sources, then the above mentioned test on the source parameters can be used to test whether these sources differ significantly given the source hypothesis and given an initial guess of the noise variance. Therefore these tests will be called a-priori tests.

2.3 Optimization

We describe two routines to optimize the D-optimality criterion. The continuous optimizer allows for sensors around the entire scalp. The discrete optimizer chooses among predetermined positions, for example in an electrocap.

If sensors are allowed on the entire scalp, then the problem is continuous. In this case minimization of $|\mathbf{C}|$ can be carried out by standard nonlinear optimization routines. We use the quasi-Newton algorithm with finite difference gradients.

If sensors are restricted to predetermined positions, then the design problem is discrete. It consists of choosing a fixed number of N sensors out of L

candidate sensors to maximize $|\mathbf{G}'\mathbf{G}|$. A naive approach is to compute $|\mathbf{G}'\mathbf{G}|$ for $L!/N!(L-N)!$ configurations, and select the optimal configuration. However, even a limited task of selecting $N=30$ out of $L=60$ sensors, leads to huge computational difficulties: calculation and comparison of about 10^{16} configurations. Therefore a more efficient technique is required, for example the exchange algorithm developed by Fedorov [1,3].

In the exchange algorithm a sensor from a starting design is exchanged for a candidate sensor if this exchanging improves $|\mathbf{G}'\mathbf{G}|$ to a maximal extent. This process is repeated until no further increase in the determinant can be attained. In the following it is indicated how the required computations can be carried out efficiently. If a sensor i is added to a design, then $\mathbf{G}'$ becomes $[\mathbf{G}', \mathbf{g}_i]$ and $\mathbf{G}'\mathbf{G}$ becomes $[\mathbf{G}'\mathbf{G} + \mathbf{g}_i\mathbf{g}_i']$, where $\mathbf{g}_i$ is the P vector with first order partial derivatives at sensor i . If $\mathbf{G}'\mathbf{G}$ is non-singular then

$$|\mathbf{G}'\mathbf{G} + \mathbf{g}_i\mathbf{g}_i'| = |\mathbf{G}'\mathbf{G}| (1 + \mathbf{g}_i'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_i) .$$

The added sensor i should give a maximal increase in the determinant. That is, it should maximize:

$$\frac{|\mathbf{G}'\mathbf{G}| (1 + \mathbf{g}_i'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_i)}{|\mathbf{G}'\mathbf{G}|} = 1 + \mathbf{g}_i'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_i$$

Similar relations hold for the deletion of a sensor j , in that case $1 - \mathbf{g}_j'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_j$ should be minimal. The criterion for exchanging a candidate sensor i for a design sensor j is derived analogously. A sensor j should be exchanged for a sensor i if it maximizes (where $\mathbf{D}=\mathbf{G}'\mathbf{G}$):

$$\frac{|\mathbf{D} + \mathbf{g}_i\mathbf{g}_i' - \mathbf{g}_j\mathbf{g}_j'|}{|\mathbf{D}|}$$

which can be rewritten as:

$$\begin{aligned} \frac{|\mathbf{D} + \mathbf{g}_i\mathbf{g}_i'| (1 - \mathbf{g}_j'[\mathbf{D} + \mathbf{g}_i\mathbf{g}_i']^{-1}\mathbf{g}_j)}{|\mathbf{D}|} = \\ \frac{|\mathbf{D}|(1 + \mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_i)(1 - \mathbf{g}_j'[\mathbf{D} + \mathbf{g}_i\mathbf{g}_i']^{-1}\mathbf{g}_j)}{|\mathbf{D}|} = \end{aligned}$$

$$\begin{aligned} (1 + \mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_i) \left(1 - \mathbf{g}_j' \left[\mathbf{D}^{-1} - \frac{\mathbf{D}^{-1}\mathbf{g}_i\mathbf{g}_i'\mathbf{D}^{-1}}{1 + \mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_i} \right] \mathbf{g}_j \right) = \\ 1 - \mathbf{g}_j'\mathbf{D}^{-1}\mathbf{g}_j + \frac{\mathbf{g}_j'\mathbf{D}^{-1}\mathbf{g}_i\mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_j}{1 + \mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_i} + \mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_i - \\ \mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_i\mathbf{g}_j'\mathbf{D}^{-1}\mathbf{g}_j + \frac{\mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_i\mathbf{g}_j'\mathbf{D}^{-1}\mathbf{g}_i\mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_j}{1 + \mathbf{g}_i'\mathbf{D}^{-1}\mathbf{g}_i} \end{aligned}$$

which equals [1,3]:

$$\begin{aligned} 1 + \mathbf{g}_i'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_i - \mathbf{g}_j'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_j - \\ \mathbf{g}_i'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_i\mathbf{g}_j'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_j + \mathbf{g}_j'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_i\mathbf{g}_i'[\mathbf{G}'\mathbf{G}]^{-1}\mathbf{g}_j \end{aligned}$$

In the first iteration this expression is evaluated for $N(L-N)$ exchanges and the best exchange is chosen.

This is not time-consuming, since $[\mathbf{G}'\mathbf{G}]^{-1}$ has to be calculated only once. This exchanging procedure is iterated until no further increment in the determinant is found. Since the choice of the initial starting design might influence the results, we restart the procedure with 5 randomly chosen starting designs.

3 Results

3.1 Simulations

The purpose of the simulations is: a) to get an indication of the improvements obtained by using continuously or discretely optimized designs; b) to investigate the use of a source hypothesis derived from a pilot; c) to investigate the benefits of an optimal design in case of multiple sources.

The data are either simulated on a standard design with 41 sensors uniformly distributed around the scalp, or on an optimal set of 41 sensors. Since the simulations are very time consuming, they concern only one or two sources in the computationally simple homogenous sphere with a radius $r=10.00$ cm. Normally distributed homoscedastic and uncorrelated noise with zero mean is added to the signals. The noise variance is set to a realistic value, yielding a signal to noise ratio of 7.10dB on the standard configuration. The sources are estimated by OLS. To get an indication of the precision of the estimates, the simulations are repeated 400 times.

Single source

First, we continuously optimized precision given a true source hypothesis. The optimal configuration yielded a considerable improvement (Table 1).

Table 1: *The improvement on optimal designs. The improvement is defined as the ratio of standard errors on optimal and standard designs, averaged over source parameters.*

Design	Improvement
Continuous, known source	.41
Discrete , known source	.74
Continuous, source hypothesis	.42

Second, we derived a discrete optimal design given a true source hypothesis. The discrete optimizer chose 41 sensors out of a candidate set of 105 sensors, which consisted of the aforementioned 41 standard sensors plus 64 additional sensors. In Table 1 it can be seen that the discrete optimizer improves precision, although self evidently less than the continuous optimizer.

Third, we investigated whether the method is also valid if it is based on a source hypothesis derived from a pilot. We selected, from the distribution of 400 estimates on the standard configuration, an estimate that was located at a maximal distance (0.72 cm) from the true source. Subsequently, we continuously optimized the sensor configuration for this estimated source. We then performed simulations for this optimal configuration and the original true source. In Table 1 it is shown that all results are comparable to the design computed from a true source hypothesis. This indicates that although a pilot may be sub optimal, it may guide the choice of a sensor configuration for an optimal experiment.

Two sources

If a sensor configuration does not provide enough information to separate sources, then the source estimates are very unreliable. Therefore it is essential to determine a-priori whether the design contains enough information to separate sources, and to optimize the design if this is not the case.

We applied this method to two sources separated 3.00 cm, and mirroring each other in the midline. The data were modeled by two or one source(s). If the sensor configuration does not provide enough

information to separate the sources, then the a-priori test on the y coordinate of the sources should be non-significant. Moreover, the Lack of Fit test [4] should then indicate in most of the simulation replicates that there is no lack of fit for the single source model.

The a-priori test on the y coordinate indeed indicated that source locations did not differ significantly on the standard configuration ($F=.95$, $df=1,29$, $p=.95$). This a-priori test was indicative of the simulation results: the single source did not fit the data in only 14% of the replicates. Therefore, the a-priori test reliably indicates that the standard design does not contain enough information to separate the sources. Results improve on the optimal configuration. First, the a-priori test was significant ($F=28.62$, $df=1,29$, $p<.001$). Second, the single source did not fit the data in 95% of the replicates. Therefore, the optimal sensor configuration yields enough information to separate the two sources.

3.2 Empirical illustration

Data were recorded on 64 sensors. The data were modeled by one dipole in three concentric isotropic spheres.

First we analyzed a subset of 32 approximately equally spaced sensors. This subset served as a ‘pilot’ generating a source hypothesis to compute the optimal 32 subset. We then estimated the source on the optimal subset. Finally, we estimated the source of the entire set of 64 sensors. If the proposed methodology is valid, then the standard errors should decrease on the optimal compared to the pilot subset. In addition the parameter and standard error estimates derived from 64 sensors should be approximated by the optimal subset.

Indeed, all standard errors were low on the optimal compared to the standard subset, although the standard error of ξ_1^n was higher. This may be due to the fact that the precision of all parameters is optimized, which may be at the cost of parameters that are already precise, like ξ_1^n . Moreover, the parameter estimates and standard errors of the optimal subset approximated those of the entire set of 64 sensors.

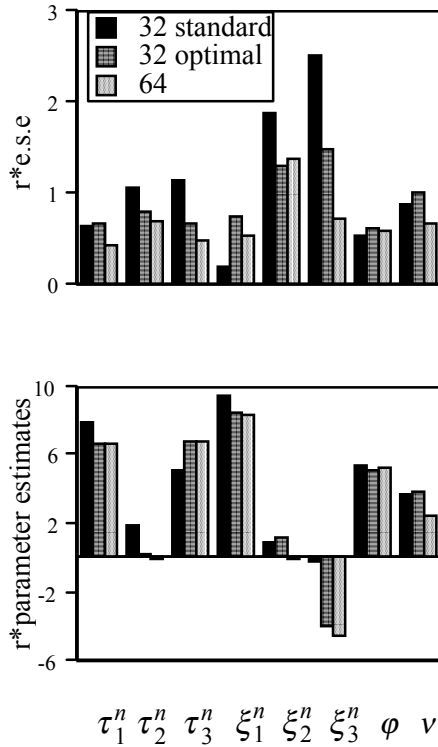


Figure 1: in the upper panel the standard error of the parameter estimates. In the lower panel the parameter estimates. All values are multiplied by r , the estimated head radius. The source is defined by respectively: 3 location direction cosines, 3 orientation direction cosines, eccentricity and strength.

4 Discussion

The method has several advantages. First, it is analytic and by consequence fast. For example, choosing 32 sensors out of a set of 64 sensors takes only a few seconds. Second, the method improves accuracy considerably. The improvement depends on the source configuration under study, so it is difficult to indicate to what extent the estimates will generally ameliorate. But to give an indication, in our simulations the standard errors halved if sensors were positioned optimally. Moreover, it became possible to separate two simultaneously active sources, which was impossible on a standard configuration. A third advantage is that experimental demands are optimized at minimal costs. Positioning sensors optimally does not cost anything. A fourth advantage is that the method may indicate a-priori that experimental requirements are not attainable. This offers the possibility to prevent experiments with insufficient power.

The method is based on the a-priori covariance matrix of the parameter estimates. Therefore this matrix should be accurate, necessitating three

requirements. First, this matrix is only accurate if the regression model can be linearized around the parameter estimates. If this is not the case, then optimal designs based on likelihood approaches may be more appropriate [2]. A second prerequisite is that noise should be homoscedastic and uncorrelated. This requirement can be avoided by incorporating a valid a-priori hypothesis on the noise covariance matrix [5]. The third requirement is an explicit source hypothesis. This hypothesis can be derived for example from previously published fMRI, PET, or EEG/MEG source analysis studies. A source hypothesis can also be derived from a pilot experiment. We have shown in simulations and in an empirical study that a pilot may give an accurate indication of experimental requirements.

Acknowledgements

We would like to thank Juha Virtanen for his help in collecting the empirical data. The research of H.M.H. has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. D.J.H. was supported by grant number 575-65-058 from the Dutch Organization for Scientific Research (NWO).

References

1. A.C. Atkinson and A.N. Donev, *Optimum Experimental Designs*. Clarendon Press, Oxford, 1992.
2. G.A.F. Seber and C.J. Wild, *Nonlinear Regression*. Wiley, New York, 1989.
3. R.C. St. John and N.R. Draper, "D-optimality for regression designs: a review", *Technometrics*, **17**, 15-23, 1975.
4. H.M. Huizenga and P.C.M. Molenaar, "Estimating and testing the sources of evoked potentials in the brain", *Multivariate Behavioral Research*, **29**, 237-262, 1994.
5. L.J. Waldorp, H.M. Huizenga, C.V. Dolan, and P.C.M. Molenaar, "Maximum likelihood electromagnetic source analysis based on a parametric noise covariance model", *submitted*.